N·H
ELSEVIER

Available at
WWW.MATHEMATICSWEB.ORG
POWERED BY SCIENCE @ DIRECT•

Journal of Computational and Applied Mathematics 154 (2003) 175–182

JOURNAL OF
COMPUTATIONAL AND
APPLIED MATHEMATICS

www.elsevier.com/locate/cam

Fixed and sequential designs for estimation in the exponential family with comparisons and applications to binomial proportions using beta priors

Mekki Terbeche, Broderick O. Oluyede*

Department of Mathematics and Computer Science, Georgia Southern University, P.O. Box 8093, Statesboro, GA 30460-8093, USA

Received 15 May 2001; received in revised form 15 September 2002

Abstract

A sequential approach to the estimation of the difference of two population means for distributions belonging to the exponential family is adopted and compared with the best fixed design. Comparisons of sequential and best fixed designs for estimation of the difference between two Bernoulli proportions with beta priors are conducted. Results on the lower bound for the Bayes risk due to estimation and expected cost are presented and shown to be of first-order efficiency. Numerical comparisons for the Bernoulli distribution with beta priors are presented.

MSC: 62L12

Keywords: Estimation; Loss function; Prior distribution; Sequential design

1. Introduction

The family of exponential type distributions play an important role in a wide variety of areas in probability and statistics. The gamma distribution which belong to the family of exponential distributions is used to model lifetimes of various practical situations including but not limited to lengths of time between catastrophic events (floods, earthquakes and so on), lengths of time between emergency arrivals at a hospital and distance traveled by a wildlife ecologist between sighting of an endangered species. The exponential distribution which is a special case of the gamma distribution

* Corresponding author.
E-mail address: boluyede@gasou.edu (B.O. Oluyede).

PII: S0377-0427(02)00820-8

have been used to describe the amount of time between occurrences of random events such as those given above. Further examples of the exponential type distributions include the Poisson and binomial distributions. In Bayesian theory, the beta distribution is a conjugate family for the Bernoulli distribution. In this paper, we consider the problem of designing an experiment to estimate the difference between two population means for distributions belonging to the exponential family plus expected cost of drawing samples from either groups using a Bayesian approach. We explore and compare the Bayes risk due to estimation plus the expected cost of sampling. Numerical comparisons are presented. These comparisons clearly indicate that random designs are more efficient.

This paper is organized as follows. Section 2 contains some preliminaries and basic results for the class of exponential type distributions including the binomial and beta distributions. Section 3 contain results on the sequential and best fixed designs. Some bounds are presented. In Section 4, we present some applications. Numerical results on the comparisons of the Bayes risk for the procedures described in Section 3 are presented in Section 5. This paper concludes with a discussion.

2. Preliminaries and basic results

In this section, we consider the family of exponential-type probability distributions on the real line, given by the family of densities $\mathscr{G}$ with respect to the Lebesgue measure. A natural form of an exponential family is as follows. Suppose C is one-to-one and $\eta = C(\theta)$, then $\theta = C^{-1}(\eta)$ and

$$f(x,\theta) = \exp\{\eta T(x) + S(x) + d(C^{-1}(\eta))\}, \tag{1}$$

where $f \in \mathscr{G}$. In this setting $E(T(X)) = -d_0'(\eta)$ and $\mathrm{Var}(T(X)) = -d_0''(\eta)$, where $d_0(\eta) = d(C^{-1}(\eta))$, $d_0'(\eta)$ and $d_0''(\eta)$ denote the first and second derivative of $d_0(\eta)$ with respect to η. In this paper we let $X_1, X_2, \ldots$ and $Y_1, Y_2, \ldots$ denote independent random variables with distributions $f_\theta(x) = \exp\{\theta x - \psi(\theta)\}$ and $g_\omega(x) = \exp\{\omega x - \varphi(\omega)\}$, respectively. Then $E_\theta(X) = \psi'(\theta)$ and $E_\omega(X) = \varphi'(\omega)$.

Suppose that a total of t observations are available, our objective is to estimate $\psi'(\theta) - \varphi'(\omega)$ with square error loss. In Bayes theory, the unknown parameter is a random variable. Let $\theta \in \Theta$ and $\pi(\theta)$ be the prior density of θ. The prior distribution expresses the state of knowledge or ignorance about the parameter θ before the analysis of the sample data. The a priori uncertainty is modeled by the use of a prior distribution which represents all that is known or assumed about the unknown parameters. In this note we consider conjugate prior distributions. The conditional distribution of θ given the sample denoted by $\pi(\theta|X)$ is called the posterior density of θ given X. The conditional density of X given θ is the likelihood function once X is observed.

Definition 1. Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed (i.i.d.) with probability density function (p.d.f.) $f(x|\theta)$. If the prior distribution $\pi(\theta)$ and the posterior distribution $\pi(\theta|X)$ of θ belong to the same family, then this family of distribution for θ is referred to as a conjugate family.

The beta distribution is a conjugate prior distribution for the binomial sampling model. The use of the beta distribution is recommended on the basis of mathematical tractability as guaranteed by the natural conjugate criterion as well as its versatility. Weiler [2] showed that the effect of assuming a beta distribution when the true distribution is not of the beta type, is negligible in many practical situations and applications. Severe deviations in the beta prior parameters values produce only slight changes in the posterior distributions. This in itself accounts for why the beta distribution is the most widely used prior distribution.

Definition 2. The Bayes risk of an estimate $\hat{\delta}$ with respect to the prior distribution $\pi(\theta)$ is

$$r(\theta,\hat{\delta}) = E[R(\theta,\hat{\delta})], \tag{2}$$

where $R(\theta,\hat{\delta}) = E[L(\theta,\hat{\delta})]$ and $L(\theta,\hat{\delta})$ is the loss function.

Now let m and n denote the number of observations sampled from the two populations and $t=m+n$ the total number of observations. We let $F_t = \sigma(X_1,\ldots,X_m,Y_1,\ldots,Y_n)$ be the σ-algebra generated by $X_1,\ldots,X_m$ and $Y_1,\ldots,Y_n$. The Bayes risk of a design Δ is given by

$$R(\Delta) = E[\mathrm{Var}(\psi'(\theta)|F_t) + \mathrm{Var}(\varphi'(\omega)|F_t)], \tag{3}$$

where Var(.) denote the variance of the argument in parenthesis.

3. Sequential and best fixed designs

In this section, we compare the best fixed design with the sequential optimal random design. Let c_1 and c_2 be the cost of sampling per observation from populations 1 and 2, respectively.

The Bayes risk due to estimation plus expected sampling cost is given by

$$\begin{aligned} R(\Delta) &= E[\mathrm{Var}(\psi'(\theta) - \varphi'(\omega))|F_t + c_1 m + c_2 n] \\ &= E\left[\frac{U_m}{m+r} + \frac{V_n}{n+s} + c_1 m + c_2 n\right], \end{aligned} \tag{4}$$

where $U_m = E[\psi''(\theta)|F_t]$ and $V_n = E[\varphi''(\omega)|F_t]$, r and s are fixed and depend on the posteriors, m and n are unknown. The objective or goal is to minimize $R(\Delta)$.

Theorem 1. *Let c_1 and c_2 be such that $c_1/(c_1+c_2) \to \gamma_1$, as c_1, $c_2 \to 0$, $0 < \gamma_1 < 1$ and $\gamma_2 = 1-\gamma_1$. Then for any random design Δ,*

$$\liminf_{c_1,c_2\to 0}\left(\frac{R(\Delta)}{(c_1+c_2)^{1/2}}\right) \geqslant 2E[(\gamma_1\psi''(\theta))^{1/2} + (\gamma_2\varphi''(\omega))^{1/2}].$$

Proof. See Terbeche [1]. □

In the sequential allocation, for a fixed total number of observations the problem is to allocate the number of observations to be taken from each population to achieve or nearly achieve some optimality condition such as minimizing the Bayes risk when the allocation is done sequentially. That is, at each stage the decision to observe X or Y may depend on available information from all previous stages. The fully sequential design achieves the lower bound. That is,

$$\liminf_{c_1,c_2\to 0}\left(\frac{R(\Delta)}{(c_1+c_2)^{1/2}}\right)=2E[(\gamma_1\psi''(\theta))^{1/2}+(\gamma_2\varphi''(\omega))^{1/2}]. \tag{5}$$

In the best fixed design or policy the risk function $R(\Delta)$ is minimized as a function of fixed sample sizes m and n and is denoted by $(m^{\mathscr{F}},n^{\mathscr{F}},k^{\mathscr{F}})$, where $m^{\mathscr{F}}+n^{\mathscr{F}}=k^{\mathscr{F}}$. This policy is asymptotically the best among the nonsequential or nonrandom policies. The best fixed design is determined by $m^{\mathscr{F}}=(E[\psi''(\theta)]/c_1)^{1/2}$ and $n^{\mathscr{F}}=(E[\varphi''(\omega)]/c_2)^{1/2}$, and achieves the lower bound, under some regularity conditions.

Theorem 2. *Let $\Delta_{\mathscr{S}}$ and $\Delta_{\mathscr{F}}$ denote the first order sequential and fixed designs respectively. Then*

$$0\leqslant\liminf_{c_1,c_2\to 0}\frac{R(\Delta_{\mathscr{S}})}{R(\Delta_{\mathscr{F}})}\leqslant 1. \tag{6}$$

Proof. The proof follows from Jensen's inequality. □

Theorem 3. *If $c_1=c_2$ then*

$$\liminf_{c_1,c_2\to 0}\frac{R(\Delta_{\mathscr{S}})}{R(\Delta_{\mathscr{F}})}=\frac{E(\psi''(\theta))^{1/2}+E(\varphi''(\omega))^{1/2}}{(E\psi''(\theta))^{1/2}+(E\varphi''(\omega))^{1/2}}. \tag{7}$$

Proof. We have

$$\liminf_{c_1,c_2\to 0}\frac{R(\Delta_{\mathscr{S}})}{R(\Delta_{\mathscr{F}})}=\frac{(\gamma_1)^{1/2}E(\psi''(\theta))^{1/2}+(\gamma_2)^{1/2}E(\varphi''(\omega))^{1/2}}{(\gamma_1E\psi''(\theta))^{1/2}+(\gamma_2E\varphi''(\omega))^{1/2}}. \tag{8}$$

If $c_1=c_2$, then $\gamma_1=\gamma_2$ and the result follows. □

Corollary 1. *If the sampling costs c_1 and c_2 are equal, and $\psi''(\theta)=\varphi''(\omega)$, then*

$$\liminf_{c_1,c_2\to 0}\frac{R(\Delta_{\mathscr{S}})}{R(\Delta_{\mathscr{F}})}=\frac{E(\psi''(\theta))^{1/2}}{(E\varphi''(\omega))^{1/2}}. \tag{9}$$

Proof. The proof follows immediately from Theorem 3. □

4. Applications

Now let the distribution of the random variables X and Y be given by $f(x,\theta)$ and $g(y,\omega)$ respectively, where

$$f(x,\theta)=\theta^x(1-\theta)^{1-x}, \tag{10}$$

$x=0,1,\ 0<\theta<1$ and

$$g(y,\omega)=\omega^y(1-\omega)^{1-y}, \tag{11}$$

$y=0,1,\ 0<\omega<1$. We assume that θ and ω are independent and distributed as Beta(a,b), $a>0$, $b>0$ and Beta(c,d), $c>0$, $d>0$. It follows therefore from Theorem 1 that

$$\liminf_{c_1,c_2\to 0}\frac{R(\Delta_{\mathscr{S}})}{R(\Delta_{\mathscr{F}})}=\frac{E(\theta(1-\theta))^{1/2}+E(\omega(1-\omega))^{1/2}}{(E[\theta(1-\theta)])^{1/2}+(E[\omega(1-\omega)])^{1/2}}, \tag{12}$$

where

$$E(\theta(1-\theta))^{1/2}=\frac{\Gamma(a+1/2)\Gamma(b+1/2)}{(a+b)\Gamma(a)\Gamma(b)}, \tag{13}$$

$a>0,\ b>0,$

$$E(\omega(1-\omega))^{1/2}=\frac{\Gamma(c+1/2)\Gamma(d+1/2)}{(c+d)\Gamma(c)\Gamma(d)}, \tag{14}$$

$c>0,\ d>0$. Similarly,

$$E[\theta(1-\theta)]=ab/(a+b+1)(a+b), \tag{15}$$

$a>0,\ b>0$, and

$$E[\omega(1-\omega)]=cd/(c+d+1)(c+d), \tag{16}$$

$c>0,\ d>0$. For the beta distribution, that is, $\theta\sim Beta(a,b)$, it is well known that

$$E(\theta)=\frac{a}{a+b}$$

and

$$\mathrm{Var}(\theta)=\frac{ab}{(a+b+1)(a+b)^2}.$$

Similarly, if $\omega\sim Beta(c,d)$, then

$$E(\omega)=\frac{c}{c+d}$$

and

$$\mathrm{Var}(\omega)=\frac{cd}{(c+d+1)(c+d)^2}.$$

The ratio of the sequential to the best fixed design is

$$\frac{R(\Delta_{\mathscr{S}})}{R(\Delta_{\mathscr{F}})}=\frac{\frac{\Gamma(a+1/2)\Gamma(b+1/2)}{(a+b)\Gamma(a)\Gamma(b)}+\frac{\Gamma(c+1/2)\Gamma(d+1/2)}{(c+d)\Gamma(c)\Gamma(d)}}{(ab/(a+b+1)(a+b))^{1/2}+(cd/(c+d+1)(c+d))^{1/2}}, \tag{17}$$

$a>0,\ b>0,\ c>0,\ d>0.$

Theorem 4. *(a) If $a=c$ and $b=d$, then*

$$\liminf_{c_1,c_2\to 0}\frac{R(\Delta_{\mathscr{S}})}{R(\Delta_{\mathscr{F}})}=\frac{\Gamma(a+1/2)\Gamma(b+1/2)}{a^{1/2}\Gamma(a)b^{1/2}\Gamma(b)}. \tag{18}$$

(*b*) *For any fixed b*

$$\frac{R(\Delta_{\mathcal{S}})}{R(\Delta_{\mathcal{F}})} \to \frac{\Gamma(b+1/2)}{b^{1/2}\Gamma(b)}, \tag{19}$$

as $a \to \infty$, *and as* $a, b \to \infty$

$$\frac{R(\Delta_{\mathcal{S}})}{R(\Delta_{\mathcal{F}})} \to 1. \tag{20}$$

(*c*) *If* $a = b = c = d$, *then*

$$\frac{R(\Delta_{\mathcal{S}})}{R(\Delta_{\mathcal{F}})} = \left(\frac{\Gamma(a+1/2)}{a^{1/2}\Gamma(a)}\right)^2 \left(\frac{2a+1}{2a}\right)^{1/2}. \tag{21}$$

(*d*) *If* $a = b = c = d = 1$, *then*

$$\frac{R(\Delta_{\mathcal{S}})}{R(\Delta_{\mathcal{F}})} = [\Gamma(3/2)]^2(3/2)^{1/2} = 0.9619. \tag{22}$$

(*e*) *If* $a, b \to 0$ *then*

$$\frac{R(\Delta_{\mathcal{S}})}{R(\Delta_{\mathcal{F}})} \to 0. \tag{23}$$

5. Numerical comparisons

In this section we examine the ratio of the sequential to the best fixed designs for the estimation problem. We consider the case of balanced and unbalanced designs. This numerical study is conducted for the case of Bernoulli proportions with beta priors (see Tables 1–4).

6. Discussion

We have established that the sequential procedure for the problem of estimating the difference of the means of two independent populations from the exponential family with conjugate priors when

Table 1
Relative efficiency for balanced design

a/b	10^{-10}	0.0010	0.0100	0.1000	1	10	50	100
10^{-10}	$2 \cdot 10^{-5}$	$3 \cdot 10^{-5}$	$3 \cdot 10^{-5}$	$2 \cdot 10^{-5}$	10^{-5}	10^{-5}	10^{-5}	10^{-5}
0.0010	$3 \cdot 10^{-5}$	0.0701	0.0938	0.0915	0.0701	0.0580	0.0564	0.0562
0.0100	$3 \cdot 10^{-5}$	0.0938	0.2183	0.2749	0.2186	0.1811	0.1761	0.1755
0.1000	$2 \cdot 10^{-5}$	0.0915	0.2749	0.6002	0.6061	0.5125	0.4987	0.4969
1	10^{-5}	0.0701	0.2186	0.6061	0.9619	0.9141	0.8926	0.8895
10	10^{-5}	0.0580	0.1811	0.5125	0.9141	0.9994	0.9933	0.9908
50	10^{-5}	0.0564	0.1761	0.4987	0.8926	0.9933	0.9999	0.9996
100	10^{-5}	0.0562	0.1755	0.4969	0.8895	0.9908	0.9996	1.0000

Table 2
Relative efficiency for fixed means at $\frac{10}{11}$

a/b	10^{-10}	0.0010	0.0100	0.1000	1	10	50	100
10^{-10}	10^{-5}	0.0299	0.0938	0.2749	0.6061	0.9141	0.9817	0.9954
0.0010	0.0299	0.0299	0.0784	0.2516	0.5809	0.8846	0.9509	0.9599
0.0100	0.0938	0.0784	0.0938	0.2298	0.5416	0.8335	0.8972	0.9057
0.1000	0.2749	0.2516	0.2298	0.2749	0.5057	0.7560	0.8113	0.8188
1	0.6061	0.5809	0.5416	0.5057	0.6061	0.7815	0.8232	0.8289
10	0.9141	0.8846	0.8335	0.7560	0.7815	0.9141	0.9485	0.9532
50	0.9817	0.9509	0.8972	0.8113	0.8232	0.9485	0.9817	0.9863
100	0.9954	0.9599	0.9057	0.8188	0.8289	0.9532	0.9863	0.9908

Note: With prior means fixed at $\frac{10}{11}$, $a = 10b$ and $c = 10d$.

Table 3
Relative efficiency for $\mu_1 = \mu_2$ and $\sigma_1 < \sigma_2$

(a, c)	Ratio
$(10^{-10}, 10^{-11})$	0.0000
$(0.001, 0.0001)$	0.0676
$(0.01, 0.001)$	0.2089
$(0.1, 0.01)$	0.5462
$(1, 0.1)$	0.8512
$(10, 1)$	0.9809
$(50, 10)$	0.9986
$(100, 50)$	0.9997

Table 4
Relative efficiency for $\mu_1 < \mu_2$ and $\sigma_1 = \sigma_2$

(a, c)	Ratio
$(10^{-10}, 10^{-9})$	0.0000
$(0.001, 0.01)$	0.0938
$(0.01, 0.1)$	0.2749
$(0.1, 1)$	0.6061
$(1, 10)$	0.9141
$(10, 50)$	0.9933
$(50, 100)$	0.9996
$(100, 120)$	1.0000

Note: $b = c$, $a = d$ and $a < b$.

compared with the best fixed design reveals the superiority of the random design. The lower bounds for the Bayes risk plus the expected costs is achieved. Application of the results to the binomial distribution using beta priors as well as numerical comparisons are conducted. For the balanced designs $E(\theta) = E(\omega)$ and $\mathrm{Var}(\theta) = \mathrm{Var}(\omega)$, that is $a = c$ and $b = d$. The ratio of the sequential to the best fixed design $R(\Delta_{\mathscr{S}})/R(\Delta_{\mathscr{F}})$ is given by Eq. (17). This ratio is symmetric in a, b, c and d. It is computed for values of a, b, c and d and given in the tables in Section 5. There are other random designs that are of interest including the two stage design, myopic design. See Terbeche [1]. These designs seem to perform better than the best fixed design.

Acknowledgements

The authors wish to thank the referees for constructive criticisms which lead to substantial improvement of the paper.

References

[1] M. Terbeche, Sequential design for estimation, Thesis, Florida Institute of Technology, 2000.
[2] H. Weiler, The use of incomplete beta functions for prior distributions in binomial sampling, Technometrics 7 (1965) 335–347.